

Scott Sanner
Marcus Hutter (Eds.)

LNAI 7188

Recent Advances in Reinforcement Learning

9th European Workshop, EWRL 2011
Athens, Greece, September 2011
Revised Selected Papers

Lecture Notes in Artificial Intelligence 7188

Subseries of Lecture Notes in Computer Science

LNAI Series Editors

Randy Goebel

University of Alberta, Edmonton, Canada

Yuzuru Tanaka

Hokkaido University, Sapporo, Japan

Wolfgang Wahlster

DFKI and Saarland University, Saarbrücken, Germany

LNAI Founding Series Editor

Joerg Siekmann

DFKI and Saarland University, Saarbrücken, Germany

Scott Sanner Marcus Hutter (Eds.)

Recent Advances in Reinforcement Learning

9th European Workshop, EWRL 2011
Athens, Greece, September 9-11, 2011
Revised Selected Papers

 Springer

Series Editors

Randy Goebel, University of Alberta, Edmonton, Canada
Jörg Siekmann, University of Saarland, Saarbrücken, Germany
Wolfgang Wahlster, DFKI and University of Saarland, Saarbrücken, Germany

Volume Editors

Scott Sanner
NICTA and the Australian National University
7 London Circuit
Canberra, ACT 2601 Australia
E-mail: Scott.Sanner@nicta.com.au

Marcus Hutter
Australian National University
Research School of Computer Science
Canberra, ACT 0200 Australia
E-mail: marcus.hutter@anu.edu.au

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-29945-2 e-ISSN 978-3-642-29946-9
DOI 10.1007/978-3-642-29946-9
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012937733

CR Subject Classification (1998): I.2.6, I.2, F.1-2, G.3, C.1.3

LNCS Sublibrary: SL 7 – Artificial Intelligence

© Springer-Verlag Berlin Heidelberg 2012

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This volume contains the papers presented at EWRL 2011: the 9th European Workshop on Reinforcement Learning held in Athens Greece, September 9-11, 2011. The workshop was co-located with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2011). The technical program of EWRL 2011 contained 38 plenary talks by authors of submitted papers and 4 plenary invited talks:

- Peter Auer (University of Leoben, Leoben, Austria), “UCRL and Autonomous Exploration”;
- Kristian Kersting (Fraunhofer IAIS and University of Bonn, Bonn, Germany), “Increasing Representational Power and Scaling Inference in Reinforcement Learning”;
- Peter Stone (University of Texas Austin, USA), “PRISM Practical RL: Representation, Interaction, Synthesis, and Mortality”;
- Csaba Szepesvári (University of Alberta, Canada), “Towards Robust Reinforcement Learning Algorithms”.

All authors of presented papers were invited to present their papers at an evening poster session held on September 10; posters were also solicited for latebreaking results with the option to submit a full paper after the workshop.

In total, there were 40 submissions to the workshop from a total of 58 authors representing 18 countries (9 European). All papers were reviewed by at least 2 reviewers, with some papers receiving a 3rd review. We had an international program committee of 24 members and 9 additional reviewers residing in 11 countries (7 European). Reviewing proceeded in two rounds: 24 of 39 submitted papers were accepted to the LNAI proceedings in the first round; after the workshop authors were allowed to resubmit previous submissions or submit new papers from latebreaking posters in a second round of reviewing in which 4 of 14 submissions (13 resubmissions, 1 new submission) were accepted to the LNAI proceedings. The present volume contains the texts of the 28 papers accepted to the LNAI proceedings in the first and second round.

The best paper prize of 500 Euro was awarded to the paper “Automatic discovery of ranking formulas for playing with multi-armed bandits” by authors Francis Maes, Louis Wehenkel and Damien Ernst (University of Liège, Belgium).

EWRL was first held in 1995 and has been held every 2-3 years since that time. Over the years, EWRL has become one of the premiere events for the discussion and dissemination of recent research results in the field of reinforcement learning. This year’s workshop was organized by Marcus Hutter (Australian National University, Canberra, Australia), Matthew Robards (Australian National University, Canberra, Australia), Scott Sanner (NICTA and Australian National University, Canberra, Australia), Peter Sunehag (Australian National University, Canberra, Australia), and Marco Weiring (University Of Groningen, Groningen, Netherlands).

We would like to thank our financial sponsors: the Artificial Intelligence Journal, the Australian National University, NICTA, and PASCAL2. Their generous

support has allowed us to continue the tradition of holding EWRL as a free event open to all interested participants. We would like to thank ECML PKDD 2011 for allowing us to co-locate with them and for advertising EWRL on their website. We are grateful that we could use the excellent conference management system EasyChair, which has been developed mainly by Andrei Voronkov and hosted at the University of Manchester; the system is cost-free. We thank Springer for their support in preparing and publishing this volume of Lecture Notes in Artificial Intelligence.

Last but not least, we wish to thank the organizers, program committee members, additional reviewers, authors, speakers and all attendees for this engaging and highly successful installment of EWRL.

February 2012

Scott Sanner
Marcus Hutter

With thanks to our EWRL 2011 sponsors:



Australian
National
University



Table of Contents

Invited Talk Abstracts

UCRL and Autonomous Exploration	1
<i>Peter Auer</i>	
Increasing Representational Power and Scaling Inference in Reinforcement Learning	2
<i>Kristian Kersting</i>	
PRISM – Practical RL: Representation, Interaction, Synthesis, and Mortality	3
<i>Peter Stone</i>	
Towards Robust Reinforcement Learning Algorithms	4
<i>Csaba Szepesvari</i>	

Online Reinforcement Learning

Automatic discovery of ranking formulas for playing with multi-armed bandits	5
<i>Francis Maes, Louis Wehenkel and Damien Ernst</i>	
Goal-Directed Online Learning of Predictive Models	17
<i>Sylvie Ong, Yuri Grinberg and Joelle Pineau</i>	
Gradient Based Algorithms with Loss Functions and Kernels For Improved On-Policy Control	29
<i>Matthew Robards and Peter Sunehag</i>	

Learning and Exploring MDPs

Active Learning of MDP models	41
<i>Mauricio Araya-López, Olivier Buffet, Vincent Thomas and François Chapillet</i>	
Handling Ambiguous Effects in Action Learning	53
<i>Boris Lesner and Bruno Zanuttini</i>	
Feature Reinforcement Learning in Practice	65
<i>Phuong Nguyen, Peter Sunehag and Marcus Hutter</i>	

Function Approximation Methods for Reinforcement Learning

Reinforcement learning with a bilinear Q function	77
<i>Charles Elkan</i>	

L1-penalized projected Bellman residual	88
<i>Matthieu Geist and Bruno Scherrer</i>	
Regularized Least Squares Temporal Difference learning with nested l2 and l1 penalization	100
<i>Matthew Hoffman, Alessandro Lazaric, Mohammad Ghavamzadeh and Remi Munos</i>	
Recursive Least-Squares Learning with Eligibility Traces	112
<i>Bruno Scherrer and Matthieu Geist</i>	
Value Function Approximation through Sparse Bayesian Modeling	124
<i>Nikolaos Tziortziotis and Konstantinos Blekas</i>	
Macro-actions in Reinforcement Learning	
Automatic construction of temporally extended actions for MDPs using bisimulation metrics	136
<i>Pablo Castro and Doina Precup</i>	
Unified Inter and Intra Options Learning Using Policy Gradient Methods	148
<i>Kfir Levy and Nahum Shimkin</i>	
Options With Exceptions	160
<i>Munu Sai and Balaraman Ravindran</i>	
Policy Search and Bounds	
Robust Bayesian reinforcement learning through tight lower bounds	172
<i>Christos Dimitrakakis</i>	
Optimizing look-ahead tree search policies	184
<i>Francis Maes, Louis Wehenkel and Damien Ernst</i>	
A Framework for Computing Bounds for the Return of a Policy	196
<i>Cosmin Paduraru, Doina Precup and Joelle Pineau</i>	
Multi-task and Transfer Reinforcement Learning	
Transferring Evolved Reservoir Features in Reinforcement Learning Tasks	208
<i>Kyriakos Chatzidimitriou, Ioannis Partalas, Pericles Mitkas and Ioannis Vlahavas</i>	
Transfer Learning via Multiple Inter-Task Mappings	220
<i>Anestis Fachantidis, Ioannis Partalas, Matthew Taylor and Ioannis Vlahavas</i>	
Multi-Task Reinforcement Learning: Shaping and Feature Selection	232
<i>Matthijs Snel and Shimon Whiteson</i>	

Multi-agent Reinforcement Learning

- Transfer Learning in Multi-agent Reinforcement Learning Domains 244
Georgios Boutsioukis, Ioannis Partalas and Ioannis Vlahavas
- An extension of a hierarchical reinforcement learning algorithm for
multiagent settings 256
Ioannis Lambrou, Vassilis Vassiliades and Chris Christodoulou

Apprenticeship and Inverse Reinforcement Learning

- Bayesian multitask inverse reinforcement learning 268
Christos Dimitrakakis and Constantin Rothkopf
- Batch, Off-policy and Model-free Apprenticeship Learning 280
Edouard Klein, Matthieu Geist and Olivier Pietquin

Real-world Reinforcement Learning

- Introduction of Fixed Mode States into Online Profit Sharing and Its
Application to Waist Trajectory Generation of Biped Robot 292
Seiya Kuroda, Kazuteru Miyazaki and Hiroaki Kobayashi
- MapReduce for Parallel Reinforcement Learning 304
Yuxi Li and Dale Schuurmans
- Compound Reinforcement Learning: Theory and An Application to
Finance 316
Tohgoroh Matsui, Takashi Goto, Kiyoshi Izumi and Yu Chen
- Proposal and Evaluation of the Active Course Classification Support
System with Exploitation-oriented Learning 328
Kazuteru Miyazaki and Masaaki Ida

Organizing Committee

Marcus Hutter	<i>(General Workshop Chair)</i>
Matthew Robards	<i>(Local Organizing Chair)</i>
Scott Sanner	<i>(Program Committee Chair)</i>
Peter Sunehag	<i>(Treasurer)</i>
Marco Wiering	<i>(Miscellaneous)</i>

Program Committee

Edwin Bonilla	NICTA and Australian National University
Emma Brunskill	University of California, Berkeley
Peter Dayan	University College London
Carlos Diuk	Princeton University
Alan Fern	Oregon State University
Fernando Fernandez	Universidad Carlos III de Madrid
Mohammad Ghavamzadeh	INRIA Lille
Marcus Hutter	Australian National University and NICTA
Kristian Kersting	Fraunhofer IAIS and University of Bonn
Shie Mannor	Technion
Ronald Ortner	University of Leoben
Joelle Pineau	McGill University
Doina Precup	McGill University
Matthew Robards	Australian National University and NICTA
Scott Sanner	NICTA and Australian National University
Guy Shani	Ben Gurion University
David Silver	University College London
Peter Sunehag	Australian National University
Prasad Tadepalli	Oregon State University
Matthew Taylor	Lafayette College
William Uther	NICTA and the University of New South Wales
Martijn Van Otterlo	Katholieke Universiteit Leuven
Thomas Walsh	University of Arizona
Marco Wiering	University of Groningen

Additional Reviewers

Bou Ammar, Haitham	Shào, Wén
Daswani, Mayank	Van Hasselt, Hado
Kalyanakrishnan, Shivaram	Vroman, Monica
Lattimore, Tor	Visentin, Daniel
Nguyen, Phuong Minh	

Author Index

A	
Araya-López, Mauricio	37
B	
Blekas, Konstantinos	120
Boutsioukis, Georgios	240
Buffet, Olivier	37
C	
Castro, Pablo	132
Charpillat, François	37
Chatzidimitriou, Kyriakos	204
Chen, Yu	312
Christodoulou, Chris	252
D	
Dimitrakakis, Christos	168, 264
E	
Elkan, Charles	73
Ernst, Damien	1, 180
F	
Fachantidis, Anestis	216
G	
Geist, Matthieu	84, 108, 276
Ghavamzadeh, Mohammad	96
Goto, Takashi	312
Grinberg, Yuri	13
H	
Hoffman, Matthew	96
Hutter, Marcus	61
I	
Ida, Masaaki	324
Izumi, Kiyoshi	312
K	
Klein, Edouard	276
Kobayashi, Hiroaki	288
Kuroda, Seiya	288
L	
Lambrou, Ioannis	252
Lazaric, Alessandro	96
Lesner, Boris	49
Levy, Kfir	144
Li, Yuxi	300
M	
Maes, Francis	1, 180

Matsui, Tohgoroh	312
Mitkas, Pericles	204
Miyazaki, Kazuteru	288, 324
Munos, Remi	96
N	
Nguyen, Phuong	61
O	
Ong, Sylvie	13
P	
Paduraru, Cosmin	192
Partalas, Ioannis	204, 216, 240
Pietquin, Olivier	276
Pineau, Joelle	13, 192
Precup, Doina	132, 192
R	
Ravindran, Balaraman	156
Robards, Matthew	25
Rothkopf, Constantin	264
S	
Sai, Munu	156
Scherrer, Bruno	84, 108
Schuurmans, Dale	300
Shimkin, Nahum	144
Snel, Matthijs	228
Snehag, Peter	25, 61
T	
Taylor, Matthew	216
Thomas, Vincent	37
Tziortziotis, Nikolaos	120
V	
Vassiliades, Vassilis	252
Vlahavas, Ioannis	204, 216, 240
W	
Wehenkel, Louis	1, 180
Whiteson, Shimon	228
Z	
Zanuttini, Bruno	49